AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 108:137–146 (1999)

Genealogy Reconstruction From Short Tandem Repeat Genotypes in an Amazonian Population

FRANCESC CALAFELL,[1] AUDREY SHUSTER,[1] WILLIAM C. SPEED,[1] JUDITH R. KIDD,[1] FRANCIS L. BLACK,[2] AND KENNETH K. KIDD[1,3]*
[1]*Department of Genetics, Yale University School of Medicine, New Haven, Connecticut 06520-8005*
[2]*Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, Connecticut 06520-8005*
[3]*Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut 06520-8005*

KEY WORDS microsatellites; Surui; parentage

ABSTRACT We have reconstructed partial genealogies in a sample of 44 SW Amazonian Rondonian Surui, in which 45 dinucleotide short tandem repeat polymorphisms had previously been typed. The genotypes of 488 pairs of individuals having an age difference of 13 or greater were compared, and parentage was excluded if a pair failed to share an allele at more than one locus. In order to test the power of this method, we computed the expected distribution of the number of exclusionary loci for such pairs of unrelated individuals, as well as that for individuals with different degrees of relatedness, and compared it to the observed distribution. We estimated that the pairs compared contained ~20% of individual pairs with a first-cousin relation or closer. A total of 25 pairs were identified as possible parent-child. In three instances, we could identify two or more children having a common parent; we computed a relatedness coefficient in order to establish whether the children were full or half sibs. The genealogies inferred show that instances of polygyny and polyandry (or, alternatively, serial mating), in addition to apparent monogamy, can be found among the Surui. The Surui sample can be used as a model for paleoanthropological populations, in which the determination of relatedness can provide further insights into the social structure of past populations. We estimate that, depending on the history of the populations and the degree of inbreeding, 10–20 highly informative nuclear loci should be typed in order to infer genealogies with acceptable confidence. Am J Phys Anthropol 108:137–146, 1999.

The advent of PCR and automated typing of short tandem repeat polymorphisms (STRPs) (also known as microsatellites) has allowed the typing of large numbers of individuals for numerous genetic markers with relative ease and speed. Thus, genotypic information for a large number of highly informative genetic markers has become available for population samples (Bowcock et al., 1994; Deka et al., 1995; Jorde et al., 1995, 1997; Pérez-Lezaun et al., 1997). This information can be used to explore the internal structure of the population sampled in terms of relatedness, genealogy, mating structure, the differential genetic contribution of some individuals to the next generation, population substructure, or recent ad-

Grant sponsor: USPHS; Grant numbers: MH39329, MH50390; Grant sponsor: NSF; Grant number: SBR-9632509; Grant sponsor: CIRIT.

Francesc Calafell's current address is Unitat d'Antropologia, Facultat de Biologia, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Catalonia, Spain.

*Correspondence to: Kenneth K. Kidd, Department of Genetics, Yale University School of Medicine, 333 Cedar St., New Haven, CT 06520-8005. E-mail: kidd@biomed.med.yale.edu

Received 2 September 1997; accepted 17 October 1998.

mixture. Intrapopulation analysis (Queller et al., 1993) based on STRPs was carried out by Morin et al. (1995) in samples of chimpanzees, and it is increasingly common in other wild populations of animals (many articles in recent issues of *Molecular Ecology* and other journals are devoted to genealogy reconstruction and relatedness in a number of species, usually for applications in conservation biology), but, to the best of our knowledge, no equivalent analysis has been undertaken in humans. Usually information gathered by field workers from the subjects sampled is sufficient to establish the relatedness among the individuals in the sample, and demographic studies can provide accurate descriptions of mating patterns, endogamy, population substructure, and migration. However, this is not always the case. Populations in which socially defined relations are markedly different from biological relations (e.g., due to frequent adoption or multiple sexual partners), may not have an accurate record of biological relations among its members. In different scenarios, samples may be available without accompanying genealogical data (e.g., archival samples), and paleoanthropological samples do not commonly display unequivocal indicators of biological relationship among individuals.

In the present paper, we demonstrate the feasibility of partial genealogy reconstruction in a traditional population (the SW Amazonian Surui), which can also serve as a model for a paleoanthropological population, in which 45 dinucleotide STRPs were typed. The Surui of Rondônia (they should not be confused with the Surui-Aikewara of Pará) constituted a population of 500 individuals around the time of fieldwork (1986) (Santos and Coimbra, 1996). They speak a language of the Tupi stock and are polygynous and patrilocal (Mindlin, 1985). The sample analyzed was from a single, 85 inhabitant village. Previous genetic studies on this Surui sample include the major histocompatibility complex (HLA) and immunoglobulin allotypes (Black, 1991; Bhatia et al., 1995), serum cholinesterases (Guerreiro et al., 1989), 31 blood groups and protein polymorphisms (Callegari-Jacques et al., 1994), and 30 nuclear restriction fragment length polymorphisms (RFLPs) (Kidd et al., 1991). They also have been included in global surveys of allele frequency variation at individual loci: DRD2 (Castiglione et al., 1995), CD4 (Tishkoff et al., 1996a), PLAT (Tishkoff et al., 1996b), and DRD4 (Chang et al., 1996).

MATERIALS AND METHODS

The Surui sample comprised 44 individuals for whom age (ranging from 2–50 years) and sex was known, and the ethnographic information collected indicated the possibility that the sample included several nuclear families. The genealogical data collected was of a social nature: the fieldworkers were informed about who a man's wives and children were but not which, if any, wife was the mother of which child or children. It was not inquired whether the children were offspring from an earlier marriage of the wife. If a son or daughter were married and had his or her own family, he or she would never be described a someone's offspring.

Forty-five STRP markers from ABI Prism Linkage Mapping Sets, panels 13–16, mapping to chromosomes 9, 10, and 11, were typed as described in Calafell et al. (1997). Allele frequencies can be retrieved from the Web site ⟨http://info.med.yale.edu/genetics/kkidd⟩. Genotype frequencies (inferred assuming codominant inheritance and no null alleles) did not show statistically significant departure from Hardy-Weinberg expectations. Mean heterozygosity was 0.5488, with a range from 0.0503–0.8103.

In the absence of complete and accurate genealogical information, the genotypes of all 488 pairs of individuals with an age difference of 13 or greater were compared, and parentage was excluded if more than one marker showed exclusionary genotypes (i.e., the pair of individuals did not share at least one allele). If exactly one incompatible locus was found, allele calling was rechecked for possible allele miscalling. If the allele calling was confirmed, a pair not sharing one allele at exactly one locus was still considered nonexclusionary, given the possibility of mutation.

In order to assess the power of the loci analyzed in positively identifying parent-child pairs, we computed the probability that a pair of unrelated individuals does not show any exclusionary autosomal geno-

types, assuming Hardy-Weinberg proportions and independence across loci, as

$$p_{ne,ur} = \prod_{l=1}^{k} (1 - p_{exl}(ur)) \tag{1}$$

where

$$p_{exl}(ur) = \sum_{i=1}^{n(l)} p_{il}^2 (1 - p_{il})^2 + 2 \sum_{\substack{i \neq j \\ i > j}}^{n(l)} p_{il} p_{jl} (1 - p_{il} p_{jl})^2 \tag{2}$$

and k is the number of loci, $n(l)$ is the number of alleles at locus l, and p_{il} and p_{jl} are the relative frequencies, respectively, of the ith and jth allele at locus l. The subscripts *ne*, *exl*, *r*, and *ur* refer to nonexclusion, exclusion, related, and unrelated, respectively. The chance that two unrelated individuals are not excluded from being parent and child (equation 1) is simply the product of the probabilities of each individual locus showing no exclusion. The chance that a locus shows exclusion (equation 2) is the probability that the two individuals are homozygotes for different alleles or that one of the two is a heterozygote and the other does not have these two alleles. Chakraborty and Jin (1993a) presented a formally different but equivalent formulation for equation 2. If the individuals in the pair are related, the probability that they do not share any allele identical by state (IBS) is

$$p_{ne,r} = \prod_{l=1}^{k} (1 - f_0(r) p_{exl}(ur)) \tag{3}$$

where $f_0(r)$ is the expected fraction of pairs presenting no alleles identical by descent for a particular degree of relatedness r. In particular, $f_0 = 0.25$ for full sibs, $f_0 = 0.5$ for half sibs or uncle-niece pairs (throughout this article, we use the phrase *uncle-niece* as shorthand for all possible gender combinations between an individual and his or her full sibs' children), and $f_0 = 0.75$ for simple first cousins (Chakraborty and Jin, 1993b).

We can extend this analysis to predict the distribution of pairs showing a particular number of exclusionary loci, following the method described by Chakraborty and Schull (1976). In particular, the expectation and variance of the number of exclusionary loci are $E(X) = \Sigma p_{ex}$ and $Var(X) = \Sigma p_{ex}(1 - p_{ex})$. The distribution of the number of exclusionary loci among all pairs of individuals in the Surui sample having an age difference 13 or greater was generated and compared to that expected under different degrees of relatedness: unrelated individuals, first cousins, half sibs, and full sibs. The expected distributions were obtained through a Monte Carlo procedure: one million pairs of individuals were generated at random and were assigned genotypes at 45 loci according to the allele frequencies observed in 45 STRP loci in the Surui. From those, and from extra simulated individuals when necessary, one million pairs each of full sibs, half sibs, and first cousins were generated. The genotypes within each pair were compared and the number of loci at which the pair did not share at least one allele (i.e., the exclusionary loci) were counted.

Next, we tested whether the groups of individuals for whom one individual had been identified as their common parent were full or half sibs. For every pair of individuals (x and y) sharing a parent, we computed the relatedness coefficient r (slightly modified from Queller and Goodnight, 1989):

$$r = \frac{\sum_{l=1}^{k} \sum_{a=1}^{2} (l_y - p^*)}{\sum_{l=1}^{k} \sum_{a=1}^{2} (1 - p^*)} \tag{4}$$

where a is an index for allelic position, p^* is the population frequency of the allele found in individual x at allelic position a, corrected by dropping individuals x and y from the sample, and I_y is an indicator variable that is set to 1 if y presents the same allele as x at position a and becomes 0 if x and y have different alleles at that position. r is an estimate of the probability that one random allele from x and another random allele from y are identical by descent. For full sibs of unrelated parents, that probability is 0.5, and for half sibs with unrelated, nonshared parents, $r = 0.25$. However, the estimation of r through equation 4 carries a relatively large standard deviation (Queller and Good-

night, 1989). We estimated the empirical distribution of *r* from 10,000 simulated pairs each of random individuals, half sibs, and full sibs, with genotypes drawn randomly with replacement with alleles at the frequencies observed in the Surui sample (Blouin et al., 1996). From those distributions, we estimated confidence intervals for *r* in unrelated individuals, half sibs, and full sibs. The empirical distributions allow for testing whether an actual *r* value falls within the range expected for half sibs or for full sibs.

RESULTS

The genotypes at 45 STRP loci of all pairs of individuals in the Surui sample having an age difference greater or equal to 13 years were compared. Thirty-five individuals were entered as possible children (i.e., there was at least one other individual among the 44 in the sample who was 13 or more years older), 15 as candidate fathers (i.e., at least one individual among the 44 in the sample was 13 or more years younger), and 12 as candidate mothers. Each individual "child" was tested against a mean of 10.03 candidate fathers (range: 3–15) and against a mean of 3.94 candidate mothers (range: 0–12). For a total of 25 of the possible children, one individual or two individuals of different sex among the possible parents showed no exclusionary loci. We did not find any case in which a child failed to show incompatible genotypes with more than two individuals or with two individuals of the same sex. Given the allele frequencies observed (Calafell et al., 1997) and from equations 1–3, the probability of not observing any exclusionary locus in a pair of unrelated individuals is 5.8×10^{-5} or 1/17,248, for a pair of simple first cousins it is 8.92×10^{-4} (1/1,121), for half sib or uncle-niece pairs it is 0.011 (1/91), and for full sib pairs it is 0.113, or $\sim$1/9. Thus, and considering the nature of the sample, we cannot exclude that some of the pairs of individuals that share at least one allele at every locus are in fact full sibs who happen to be 13 or more years apart in age. With the 21 most informative loci (Fig. 1), the probability of not excluding an unrelated individual as a parent (i.e., the type I error rate) is slightly below 0.001; with those loci, 0.7% of first-cousin pairs, 4.3% of half sibs, and 22% of full sibs would not be excluded as parent-offspring dyads (Fig. 1).

Twenty-five pairs of individuals met the criteria for being possible parent-offspring pairs: they had an age difference of 13 years or greater, and they shared an allele at a minimum of 44 of the 45 loci typed. However, as we have estimated above, full sibs have a relatively high probability of meeting the genetic criterion for parent-offspring pairs. To address this issue, we analyzed the distribution of the number of nonexclusionary loci per individual pair in order to estimate the expected number of putative parent-child pairs that have in fact other degrees of relatedness. The distribution of the number of expected exclusionary loci given the observed allele frequencies was estimated as described in Materials and Methods for unrelated individuals as well as for different degrees of relatedness: first cousins, half sibs or uncle-niece, and full sibs.

The distribution observed for pairs of individuals in the Surui sample having an age difference of 13 years or greater plus the expected distributions under different degrees of relatedness are shown in Figure 2. The observed distribution is roughly bell-shaped, with a mode at eight loci, a maximum of 19, and a secondary mode at zero exclusionary loci. The mean number of exclusionary loci observed was 8.20, with a variance of 9.98; the values expected if all the individuals were unrelated were 8.29 and 6.18. If all the pairs of individuals analyzed had the same degree of relatedness, the mode of the number of exclusionary loci would increase as the degree of relatedness decreased (Fig. 2): full sibs have a mode at two loci, half sib or uncle-niece pairs peak at four exclusionary loci, first cousins peak at six, and unrelated individuals show a mode at eight loci, as the observed distribution did. Presumably the kinship within every pair of individuals we compared varies; more remote multiple relationships would result in essentially a continuum of possible kinship values.

We used the four distributions of number of exclusionary loci with different degrees of relatedness to estimate the proportions of the different relations found within the pairs of individuals by finding the proportions

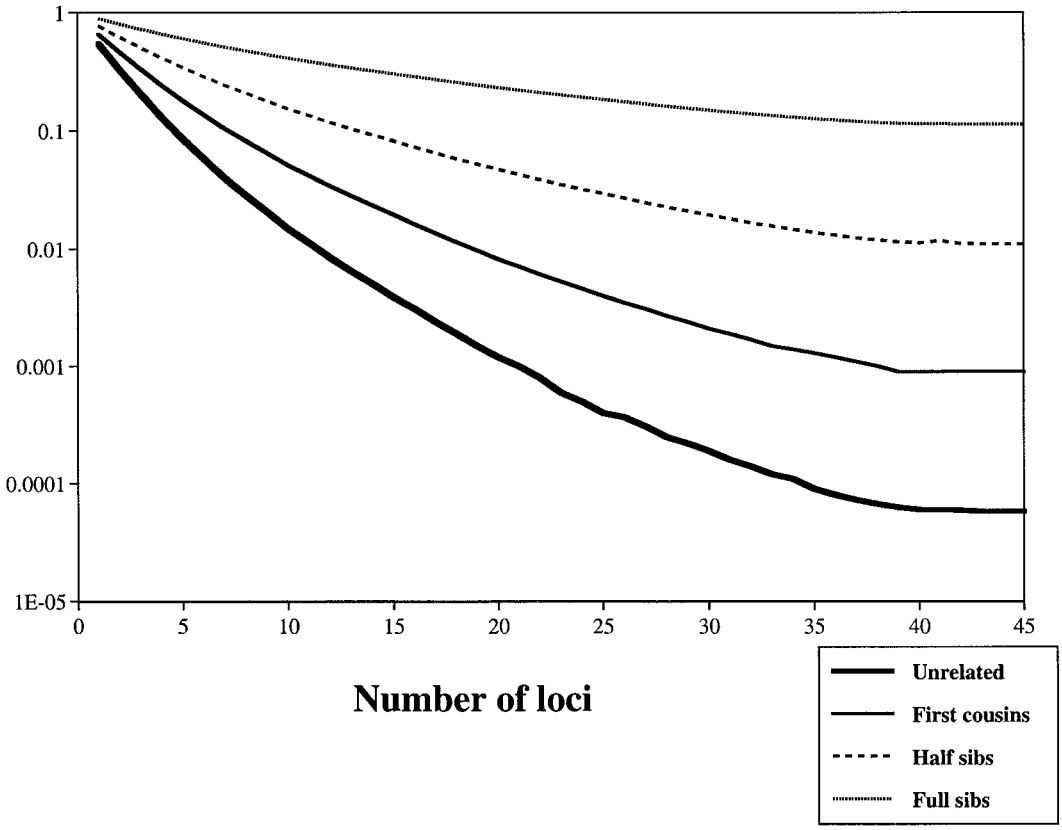

Fig. 1. Probability of not excluding a pair of Surui individuals as parent-child given different degrees of relatedness, estimated from equations 1–3, using the observed allele frequencies for the 45 STRP loci typed. Loci are sorted by decreasing informativeness (p_{ex} in equation 2).

that would fit best (by a least-square criterion) the observed distribution, conditional on the frequency of nonexclusions being equal to that observed. The best fit was obtained when the pairs in the sample comprised 79.6% unrelated pairs, 12.4% half sibs or uncle-niece, 3.4% full sibs, 4.6% parent-child pairs, and less than 0.1% first cousins. However, the mixed distribution (Fig. 2) was statistically significantly different from the observed distribution ($\chi^2 = 43.88$, 13 d.f., $P \simeq 3 \times 10^{-5}$), probably due to the heavier right tail of the observed distribution. These estimated proportions imply that the 25 nonexclusionary pairs are expected to comprise 22.45 parent-child pairs, 1.87 full sibs, 0.66 half sib or uncle-niece pairs, and 0.02 unrelated pairs. However, the fact that we imposed an age difference of at least 13 years within the pairs is expected to create a bias towards parent-child pairs and against sib pairs. Moreover, the mean age difference between nonexclusionary pairs is 27.60 ± 1.82 years, with only four cases below 20 years, and significantly larger than the age difference within exclusionary pairs (23.18 ± 0.38 years, $P = 0.006$, Mann-Whitney's U test). Therefore, in our next analyses we will treat all nonexclusionary pairs as parent-child pairs.

Ignoring for the moment redundancy of individuals in the various pairs, we found the following nonexclusionary pairs. Considering the 35 individuals young enough to have a potential parent in the sample, four had both a mother and a father, 14 had only a father, and three only a mother. Out of the 15 men old enough to be fathers, eight had children in the sample, as did five out of 12 women. Three men had six, five, and two

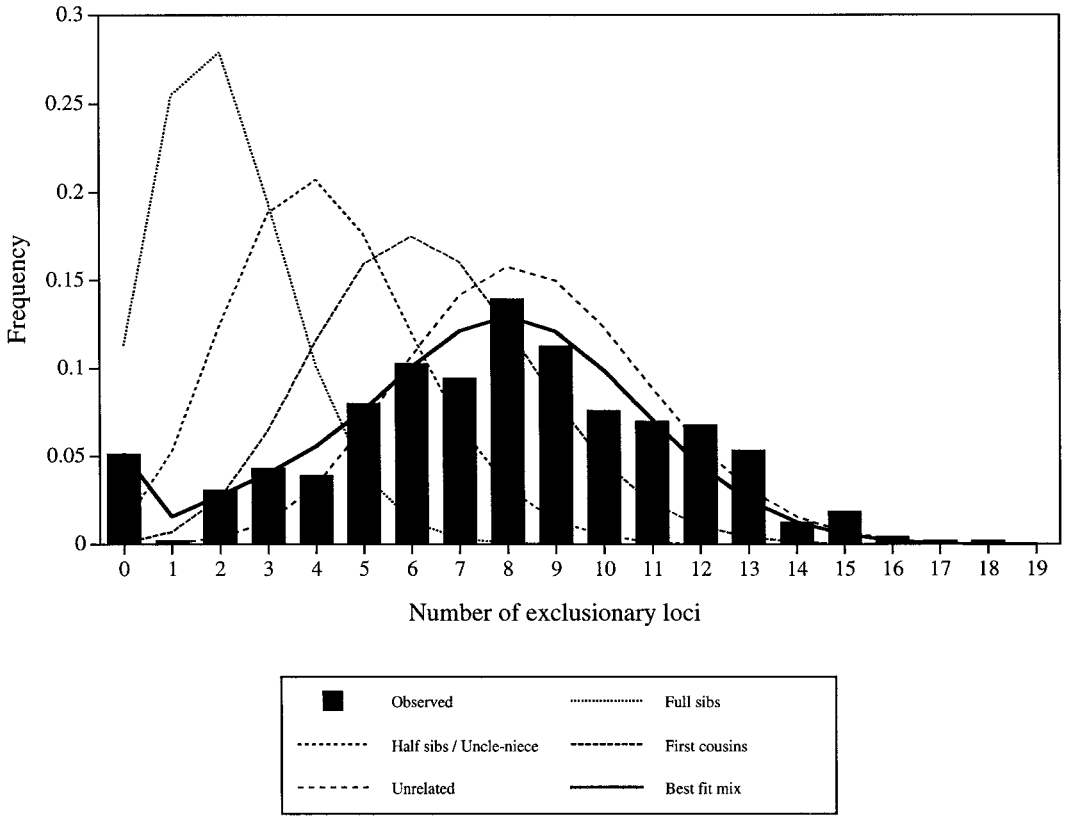

Fig. 2. Number of loci showing an exclusion in a putative parent-child pair (i.e., the two genotypes do not present any shared allele) from 45 STRP loci. Bars represent the observed number of exclusionary loci in 488 pairs of Surui individuals; lines are expected distributions if all pairs had the same degree of relatedness. The best-fit mix is a linear combination of the distributions for unrelated individuals, first cousins, half sibs, and full sibs, obtained by minimizing the squared difference with the observed distribution.

children each; one woman had three children in the sample, and five men and four women had one child each. Thus, four sibships, with six, five, three, and four sibs, respectively, were recognized; we will label them A, B, C, and D (Fig. 3). Sibship C consists of three children who shared a mother, but we identified a different putative father for each child. One of those men was also the putative father of five other children in sibship A. With that exception, we identified the putative fathers but not the mothers of children in sibships A, B, and D.

The empirical distributions of r for random pairs, half sibs, and full sibs are shown in Figure 4. We computed r relatedness coefficients for all pairs of individuals in sibships A, B, and D. For pairs of offspring from sibship A, all of whom are inferred to share the same father, r ranged from 0.120–0.319, with one exception (JK1476-JK1504). Again with that exception, we could reject with $P < 0.02$ that r was within the distribution for full sibs, but we could not reject that r was within the range for half sibs. For JK1476-JK1504, $r = 0.458$, with $P = 0.277$ for $r = 0.5$ and $P = 0.006$ for $r = 0.25$. The three individuals JK1511, JK1514, and JK1519 had among them more than two different nonpaternal alleles at five loci. The overall evidence points to JK1476 and JK1504 sharing a mother and JK1511, JK1514, and JK1519 being children of three different women. In sibship B (Fig. 3), all but one r value fell in the range 0.405–0.583, with $P < 0.05$ for belonging to the empirical distribution for half sibs and $P > 0.05$ for belonging to that of full sibs. The pair

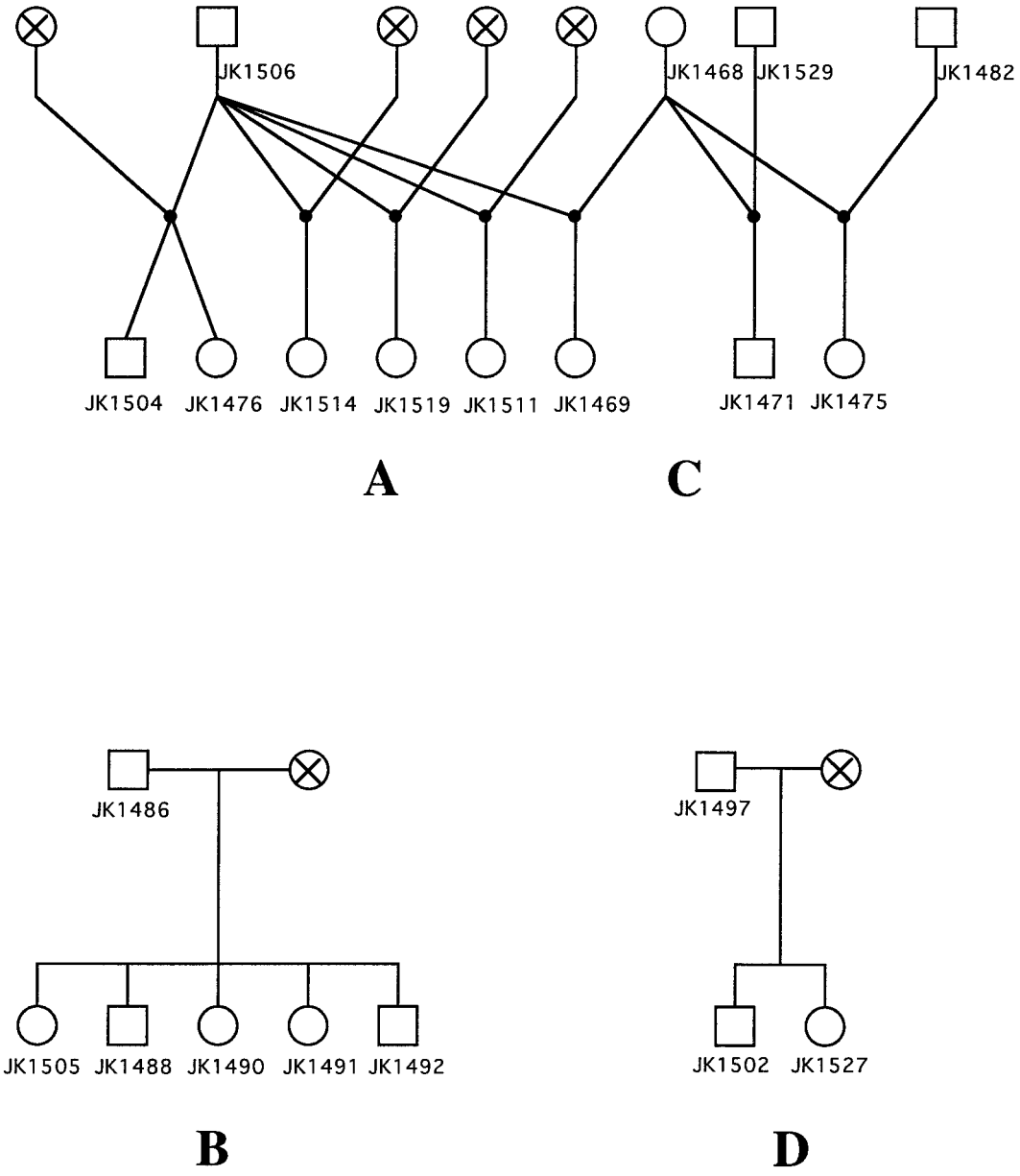

Fig. 3. Inferred relations among the Surui sample. Crossed symbols represent inferred individuals not present in the sample. Not shown: three mother-child pairs, one father-child, and one father-mother-child. Pedigree **A** refers to the children of JK1506, and pedigree **C** is defined by the children of JK1468. Pedigrees **B** and **D** refer to the only other inferred pedigrees with more than one child.

JK1492-JK1505 showed a paradoxical $r = 0.235$, which is lower than 0.5 with a nominal $P = 0.008$. Obviously, JK1492 and JK1505 cannot be half sibs to each other and share three full sibs. At all loci, we were able to reconstruct a maternal genotype from the 45 genotypes of all five sibs taken together. Thus, and considering the multiple tests done, the most likely explanation for the low relatedness coefficient between JK1492 and JK1505 is a chance effect. JK1502 and JK1527 (sibship D, Fig. 3) present $r = 0.453$,

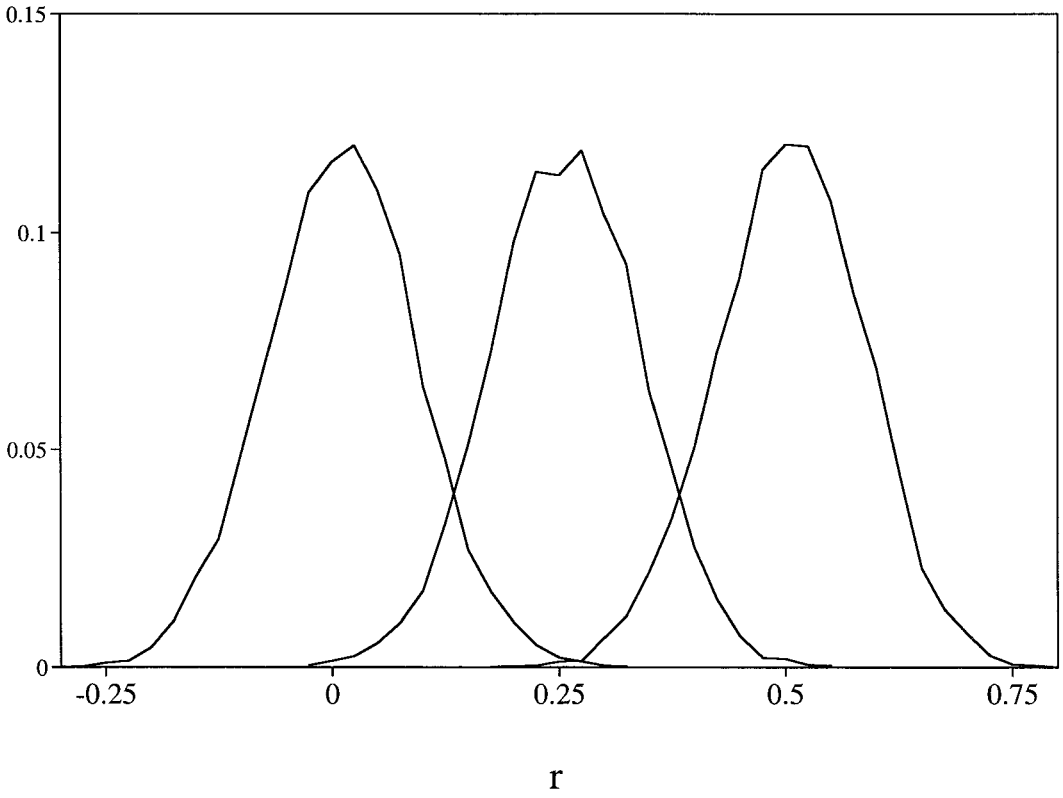

Fig. 4. Frequency distributions of the *r* relatedness coefficient (Queller and Goodnight, 1989) obtained from 10,000 simulated pairs of, from left to right, unrelated individuals, half sibs, and full sibs. Individual genotypes were simulated by drawing randomly with replacement from the allele frequencies observed in 45 loci in the Surui. Ninety-five percent empirical confidence intervals are, respectively, [−0.153, 0.178], [0.093, 0.416], and [0.340, 0.662].

which is larger than 0.25 with $P = 0.007$, and we can accept that they share a mother.

DISCUSSION

We have identified 25 parent-offspring dyads in a sample of 44 Rondônian Surui individuals, in which 45 STRP loci had been typed. Overall, we were able to place 34 of the 44 individuals in family groups. Even with the sparse genealogical information collected from the field, our genetic results were confirmed by at least 13 socially defined relationships. The same loci had also been typed in nine other populations from worldwide samples (Calafell et al., 1997). The Surui sample showed the lowest average heterozygosity among the populations typed, which results in reduced informativeness for parentage testing. Moreover, the Surui showed also the lowest heterozygosity in a survey of HLA class I antigens among 22 native South American populations (Bhatia et al., 1995). The type I error for unrelated individuals (i.e., the probability that a pair of unrelated individuals would not show exclusionary genotypes) is 439 to over 10^{10} times lower in the other populations studied for the same set of 45 loci. Using the 21 most informative loci, the type I error for unrelated individuals in the Surui is roughly 0.001; only the top seven loci would be needed in African populations in order to achieve the same error rate, (*Biaka and Mbuti Pygmies*) seven to eight in Europeans (Danes and Druze), 10–11 in East Asians (Chinese, Japanese, and Yakut), 13 in Melanesians (Nasioi), and 11 in the Maya. The low informativeness of the loci typed,

which is likely to have been caused by bottleneck events in the initial colonization of South America (Calafell et al., 1997) as well as in subsequent demographic crises, and the presence of individuals with relations other than parent-child implies that additional caution should be used in inferring genealogies. However, the large age difference observed in nonexclusionary pairs is reassuring, especially because it is significantly higher than the age difference among exclusionary pairs.

The distribution of the number of exclusionary loci among pairs of individuals with an age difference of 13 years or greater was closest to that of a sample consisting of mostly unrelated individuals, with ~20% of pairs having first cousin relatedness or closer. However, the distribution of exclusionary loci showed a right tail that was heavier than expected based on the best-fit mix of the distributions for unrelated individuals or for individuals with various degrees of relatedness. A few individuals were overrepresented in the pairs showing the highest number of exclusionary loci. Those individuals clearly clustered with the Surui when a neighbor-joining tree based on the proportion of shared alleles (computed as described by Bowcock et al., 1994) was constructed with over 500 individuals from ten worldwide populations (data not shown); thus, it is more likely that those individuals (two women) represent recent migrants from neighboring populations rather than non–Native American admixture.

Although the relatedness coefficient *r* (Queller and Goodnight, 1989) has a large variance, it takes into account information on allele frequencies, whereas other methods, based on the extent of allele sharing (Morin et al., 1995), probably do not extract all the information available (Blouin et al., 1996). Given the possible complex pattern of relatedness among the individuals, the relatively broad overlap between the confidence intervals of *r* for different degrees of relatedness, and the low informativeness of genetic markers in the Surui, we limited our analysis of relatedness to those individuals for whom one common parent had been identified and tested whether they were more likely to be half or full sibs. Using this method, we could confidently classify all sib pairs tested into half or full sibs. To the extent that our sample represents the Surui community it was drawn from, we can observe different mating patterns: a man had five children with the same woman, whereas another man had six children with five different women, one of whom had two other children with two different men. The latter cases can be attributed to polygamy and/or to serial mating. This kind of information can be incorporated into any genetic model in which effective population size is required. The difference in relatedness between full and half sibs is maximal when the nonshared parents of half sibs are unrelated to each other. However, if the unshared parents are full sibs to each other (and in fact the preferred wives for a Surui man are his sisters' daughters [Mindlin, 1985]), *r* would become 0.375. For a pair of half sibs in that situation, it would be often difficult to discern their immediate relation, and they could be easily deemed full sibs.

The genetic determination of parentage and other degrees of relatedness has an important application in archeology and paleoanthropology. The extent of the relatedness of individuals interred in the same cemetery or in the same grave can provide important clues in the reconstruction of the social structure of ancient populations. Population analysis based on ancient mtDNA has been undertaken in New World populations (Stone and Stoneking, 1993; Parr et al., 1996). However, autosomal nuclear markers are needed to establish with a reasonable level of certainty the degree of relatedness between individuals. Genotyping of STRPs from ancient DNA has been attempted with mixed results (Hauswirth et al., 1994; Ramos et al., 1995; Zierdt et al., 1996), and relatively large amounts of undegraded DNA would be required for reliable genotyping of STRPs. These restrictions would probably impose a relatively recent age for the samples that can be typed, and thus it is likely that the populations studied would be close temporally and genetically to contemporary populations. The markers typed for an intrapopulation genetic analysis should be selected carefully. Given that ancient DNA analysis and in particular ancient

nuclear DNA typing can be extremely time consuming, it is paramount to select those markers that are expected to provide the most information. Data from modern populations can be used to infer the expected informativeness of the loci and in turn can provide a framework against which to test for genetic continuity through time.

ACKNOWLEDGMENTS

The Applied Biosystems Division of Perkin Elmer provided Linkage Mapping Sets and other materials for this study, which was also supported in part by USPHS grants MH39329 and MH50390 (to K.K.K.) and NSF grant SBR-9632509 (to J.R.K.). F.C. is the recipient of a postdoctoral fellowship from CIRIT, autonomous government of Catalonia.

LITERATURE CITED

Bhatia KK, Black FL, Smith TA, Prasad ML, Koki GN. 1995. Class I HLA antigens in two long-separated populations: Melanesians and South Amerinds. Am J Phys Anthropol 97:291–305.

Black FL. 1991. Reasons for failure of genetic classifications of South Amerind populations. Hum Biol 63:763–774.

Blouin MS, Parsons M, Lacaille V, Lotz S. 1996. Use of microsatellite loci to classify individuals by relatedness. Mol Ecol 5:393–401.

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455–457.

Calafell F, Shuster A, Speed WC, Kidd JR, and Kidd KK. 1997. Short tandem repeat polymorphism in humans. Eur J Hum Genet 6:38–49.

Callegari-Jacques SM, Salzano FM, Weimer TA, Hutz MH, Black FL, Santos SEB, Guerreiro JF, Mestriner MA, Pandey JP. 1994. Further blood genetic studies on Amazonian diversity—data from Indian groups. Ann Hum Biol 21:465–481.

Castiglione CM, Deinard AS, Speed WC, Sirugo G, Rosenbaum HC, Zhang Y, Grandy DK, Grigorenko EL, Bonné-Tamir B, Pakstis AJ, Kidd JR, Kidd KK. 1995. Evolution of haplotypes at the DRD2 locus. Am J Hum Genet 57:1445–1456.

Chakraborty R, Jin L. 1993a. Determination of relatedness between individuals using DNA fingerprinting. Hum Biol 65:875–895.

Chakraborty R, Jin L. 1993b. A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. In: Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ, (editors). DNA fingerprinting: state of the science. Basel, Switzerland: Birkhäuser Verlag, pp. 153–175.

Chakraborty R, Schull WJ. 1976. A note on the distribution of the number of exclusions to be expected in paternity testing. Am J Hum Genet 28:615–618.

Chang F-M, Kidd JR, Livak KJ, Pakstis AJ, Kidd KK. 1996. The world-wide distribution of allele frequencies at the human dopamine D4 receptor locus. Hum Genet 98:91–101.

Deka R, Shriver MD, Yu LM, DeCroo S, Hundrieser J, Bunker CH, Ferrell RE, Chakraborty R. 1995. Population genetics of dinucleotide (dC-dA)n (dG-dT)n polymorphisms in world populations. Am J Hum Genet 56:461–474.

Guerreiro JF, Santos SEB, Aguiar GFS. 1989. Serum cholinesterase polymorphisms (CHE1 and CHE2 loci) among several Indian regions from Amazon region of Brazil, and segregation of the C5 variant in families. Gene Geogr 3:11–20.

Hauswirth WW, Dickel CD, Rowold DJ, Hauswirth MA. 1994. Inter- and intrapopulation studies of ancient humans. Experientia 50:585–591.

Jorde LB, Bamshad MJ, Watkins WS, Zenger R, Fraley AE, Krakowiak PA, Carpenter KD, Soodyall H, Jenkins T, Rogers AR. 1995. Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. Am J Hum Genet 57:523–538.

Jorde LB, Rogers AR, Bamshad MJ, Watkins WS, Krakowiak P, Sung S, Kere J, Harpending HC. 1997. Microsatellite diversity and the demographic history of modern humans. Proc Natl Acad Sci USA 94:3100–3103.

Kidd JR, Black FL, Weiss KM, Balazs I, Kidd KK. 1991. Studies of three Amerindian populations using nuclear DNA polymorphisms. Hum Biol 63:775–794.

Mindlin B. 1985. Nós paiter, os Surui de Rondônia. Petrópolis, Brasil: Vozes.

Morin PA, Moore JJ, Chakraborty R, Jin L, Goodall J, Woodruff DS. 1995. Kin selection, social structure, gene flow, and the evolution of chimpanzees. Science 265:1193–1201.

Parr RL, Carlyle SW, O'Rourke DH. 1996. Ancient DNA analysis of Fremont Amerindians of the Great Salt Lake wetlands. Am J Phys Anthropol 99:507–518.

Pérez-Lezaun A, Calafell F, Mateu E, Comas D, Ruiz-Pacheco R, Bertranpetit J. 1997. Microsatellite variation and the differentiation of modern humans. Hum Genet 99:1–7.

Queller DC, Goodnight KF. 1989. Estimating relatedness using genetic markers. Evolution 43:258–275.

Queller DC, Strassmann JE, Hughes CR. 1993. Microsatellites and kinship. Tree 8:285–288.

Ramos MD, Lalueza C, Girbau E, Pérez-Pérez A, Quevedo S, Turbón D, Estivill X. 1995. Amplifying dinucletode microsatellite loci from bone and tooth samples of up to 5000 years of age: more inconsistency than usefulness. Hum Genet 96:205–212.

Santos RV, Coimbra CEA Jr. 1996. Socioeconomic differentiation and body morphology in the Suruí of Southwestern Amazonia. Curr Anthropol 37:851–856.

Stone AC, Stoneking M. 1993. Ancient DNA from a pre-Columbian Amerindian population. Am J Phys Anthropol 92:463–471.

Tishkoff SA, Dietzch E, Speed W, Pakstis AJ, Kidd JR, Cheung K, Bonné-Tamir B, Santachiara-Benerecetti S, Moral P, Krings M, Pääbo S, Watson E, Risch N, Jenkins T, Kidd KK. 1996a. Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. Science 271:1380–1387.

Tishkoff SA, Ruano G, Kidd JR, Kidd KK. 1996b. Distribution and frequency of a polymorphic *Alu* insertion at the plasminogen activator locus. Hum Genet 97:759–764.

Zierdt H, Hummel S, Herrmann B. 1996. Amplification of human short tandem repeats from medieval teeth and bone samples. Hum Biol 68:185–199.